Rubric for ETL-Movies:

	Mastery **25 to > 23 points**	**Approaching Mastery** **23 to > 19 points**	**Progressing** **19 to > 16 points**	**Emerging** **16 to > 0 points**	**Incomplete**
Deliverable 1: Write an ETL function to read three data files	The ETL function does the following: ✓ The three data files are passed into the function. (10 pt) ✓ All three data sets are converted to DataFrames and the DataFrames are correct and displayed. (15 pt)	The ETL function does the following: ✓ The three data files are passed into the function. (10 pt) ✓ The Wikipedia JSON file is converted to DataFrame, and the DataFrame is displayed. (5 pt) ✓ The Kaggle metadata and MovieLens ratings data are converted to DataFrame, but the wrong DataFrames are displayed. (8 pt)	The ETL function does the following: ✓ The three data files are passed into the function. (10 pt) ✓ The Wikipedia JSON file is converted to DataFrame, but it is not displayed. (3 pt) ✓ The Kaggle metadata and MovieLens ratings data are converted to DataFrame, but the DataFrames are incorrect. (6 pt)	The ETL function does the following: ✓ The three data files are passed into the function. (10 pt) ✓ The Wikipedia JSON file is ONLY converted to a raw data file. (2 pt) ✓ The Kaggle metadata and MovieLens ratings data are converted to DataFrame, but they are not displayed. (4 pt)	No submission was received -OR- Submission was empty or blank -OR- Submission contains evidence of academic dishonesty
	Mastery **30 to > 26 points**	**Approaching Mastery** **27 to > 25 points**	**Progressing** **25 to > 20 points**	**Emerging** **20 to > 0 points**	
Deliverable 2: Extract and Transform the Wikipedia Data	✓ TV shows are filtered out and the wiki_movies DataFrame is created. (3 pt) ✓ A try-except block is used to catch errors while extracting the IMDb IDs and duplicates are dropped. (5 pt) ✓ All of the tasks for the extraction & transformation of the Wikipedia data are completed. (18 pt) ✓ The cleaned Wikipedia data is converted to a DataFrame and the DataFrame is displayed.(4 pt)	✓ TV shows are filtered out and the wiki_movies DataFrame is created. (3 pt) ✓ A try-except block is used to catch errors while extracting the IMDb IDs and duplicates are dropped. (5 pt) During the extraction & transformation of the Wikipedia data the following are done: ✓ Columns with null values are dropped. (3 pt) ✓ The non-null box office data is converted to string values. (3 pt) ✓ A regular expression code for "form_one" and "form_two" of the box office data are correct. (4 pt) ✓ THREE of the FOUR columns are cleaned. (6 pt) ✓ Wikipedia data is not cleaned but	✓ TV shows are filtered out and the wiki_movies DataFrame is created. (3 pt) ✓ A try-except block is used to catch errors while extracting the IMDb IDs and duplicates are dropped. (5 pt) During the extraction & transformation of the Wikipedia data the following are done: ✓ Columns with null values are dropped. (3 pt) ✓ The non-null box office data is converted to string values. (3 pt) ✓ A regular expression code for "form_one" and "form_two" of the box office data are correct. (4 pt) ✓ TWO of the FOUR columns are cleaned. (4 pt) ✓ Wikipedia data is not cleaned but	✓ TV shows are filtered out and the wiki_movie DataFrame is created. (3 pt) ✓ A try-except block is written, but doesn't catch errors while extracting the IMDb IDs and dropping duplicates. (3 pt) During the extraction & transformation of the Wikipedia data the following are done: ✓ Columns with null values are dropped. (3 pt) ✓ The non-null box office data is converted to string values. (3 pt) ✓ A regular expression code code for "form_one" and "form_two" of the box office data are correct. (4 pt) ✓ ONE of the FOUR columns are cleaned. (2 pt)	

		is converted to a DataFrame and displayed. (3 pt)	is converted to a DataFrame and displayed. (3 pt)	✓ Wikipedia data is not cleaned but is converted to a DataFrame and displayed. (2 pt)	
	Mastery 30 to > 27 points	**Approaching Mastery 27 to > 22 points**	**Progressing 22 to > 16 points**	**Emerging 16 to > 0 points**	
Deliverable 3: Extract and Transform the Kaggle Data	During the extraction & transformation of the Kaggle metadata the following are done: ✓ The metadata is cleaned. (4 pt) ✓ The Wikipedia and Kaggle DataFrames are merged (3 pt) ✓ The"movies" DataFrame is created and all FOUR tasks are performed. (8 pt) ✓ All THREE tasks are completed during the extraction & transformation of the MovieLens rating data. (10 pt) ✓ The Kaggle and ratings DataFrames are correct and displayed. (5 pt)	During the extraction & transformation of the Kaggle metadata the following are done: ✓ The metadata is cleaned. (4 pt) ✓ The Wikipedia and Kaggle DataFrames are merged (3 pt) ✓ The"movies" DataFrame is created but only THREE of the FOUR tasks are performed. (6 pt) ✓ All THREE tasks are completed during the extraction & transformation of the MovieLens rating data. (10 pt) ✓ The Kaggle and MovieLens rating DataFrames are displayed, but the"movies" DataFrame is incorrect. (4 pt)	During the extraction & transformation of the Kaggle metadata the following are done: ✓ The metadata is cleaned. (4 pt) ✓ The Wikipedia and Kaggle DataFrames are merged (3 pt) ✓ The"movies" DataFrame is created but only TWO of the FOUR tasks are performed. (4 pt) During the extraction & transformation of the MovieLens rating data the following are done: ✓ The ratings counts are cleaned. (3 pt) ✓ The two DataFrames are merged. (4 pt) ✓ There is an attempt to fill the empty values with “0”. (1 pt) ✓ The Kaggle and MovieLens rating DataFrames are displayed, but incorrect. (3 pt)	During the extraction & transformation of the Kaggle metadata the following are done: ✓ The metadata is cleaned. (4 pt) ✓ The Wikipedia and Kaggle DataFrames are merged but there is an error. (2 pt) ✓ The"movies" DataFrame is created but only ONE of the FOUR tasks are performed. (2 pt) During the extraction & transformation of the MovieLens rating data the following are done: ✓ The ratings counts are cleaned. (3 pt) ✓ The two DataFrames are merged but there is an error (2 pt) ✓ There is an attempt to fill the empty values with “0”. (1 pt) ✓ The Kaggle and MovieLens rating DataFrames are displayed, but incorrect. (2 pt)	
	Mastery 15 to > 14 points	**Approaching Mastery 14 to > 11 points**	**Progressing 11 to > 8 points**	**Emerging 8 to > 0 points**	
Deliverable 4: Create the Movie Database	✓ The data in the movies table in the SQL database is replaced. (5 pt) ✓ The ratings table is dropped and the MovieLens rating CSV file is added to the SQL ratings table. (5 pt) ✓ The elapsed time to add the data to the database is displayed. (5 pt)	✓ The data in the movies table in the SQL database is replaced. (5 pt) ✓ The ratings table is dropped but not all of the MovieLens rating CSV file is not added to the ratings table. (4 pt) ✓ The elapsed time to add the data to the database is displayed. (5 pt)	✓ The data in the movies table in the SQL database is not replaced. (4 pt) ✓ The ratings table is not dropped but the MovieLens rating CSV file is added to the ratings table. (3 pt) ✓ All of the elapsed time to add the data to the database is not displayed (4 pt)	✓ There is an error adding the movies table in the database. (3 pt) ✓ The ratings table is not dropped and there is an error adding the MovieLens rating CSV data. (2 pt) ✓ The elapsed time to add the data to the database is partially displayed with an error. (3 pt)	